



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

The semantics of definite expressions and the grammaticalization of THE

Citation for published version:

Gisborne, N 2012, 'The semantics of definite expressions and the grammaticalization of THE', *Studies in Language*, vol. 36, no. 3, pp. 603-644. <https://doi.org/10.1075/sl.36.3.06gis>

Digital Object Identifier (DOI):

[10.1075/sl.36.3.06gis](https://doi.org/10.1075/sl.36.3.06gis)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Studies in Language

Publisher Rights Statement:

© Gisborne, N. (2012). The semantics of definite expressions and the grammaticalization of THE. *Studies in Language*, 36(3), 603-644. 10.1075/sl.36.3.06gis

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



The semantics of definite expressions and the grammaticalization of THE*

Nikolas Gisborne

University of Edinburgh

n.gisborne@ed.ac.uk

Linguistics and English Language

School of Philosophy, Psychology and Language Sciences

The University of Edinburgh

Dugald Stewart Building

3 Charles Street

Edinburgh

EH8 9AD

United Kingdom

Abstract

This paper explores the claim that definite expressions involve universal and existential quantification from the point of view of Word Grammar, in order to establish whether the quantificational view of definiteness is compatible with a particular cognitive theory of language, and to how it compares with the familiarity treatment of definiteness. It is argued that the quantificational approach is superior to the familiarity approach in the analysis of a number of linguistic phenomena, and a number of Word Grammar analyses are presented. The paper concludes with an investigation into the grammaticalization of the English definite article, in order to compare the merits of the two approaches, and argues that the quantificational approach delivers a simpler and preferable account.

1. Introduction

We know that the English definite article, *THE*, is the result of the grammaticalization of the demonstrative pronoun *SE*. Diessel (1999) and Hawkins (2004) trace a number of grammaticalization paths that demonstrative pronouns have taken cross linguistically, and the English history is just one of the possible grammaticalization paths that demonstrative pronouns might take. It is also clear that this change is an example of grammaticalization; although both definite articles and demonstrative pronouns are grammatical words, and the main difference between them appears to be semantic, there is also a phonological difference. *THE* is clearly a clitic: it varies in form before consonants and vowels just as *A(N)* does. However, I will argue that there is no categorial change or other syntactic change and that this change involves just the loss of some

semantic content that demonstratives have but which definite articles do not have, and the phonological reduction to clitic status.

However, there are some problems. In particular, what is the right semantics of THE, or of definiteness more generally? There is a debate in the formal semantic literature about how, precisely, the English definite article ought to be treated: is it a quantifier, as Russell (1905) claimed, or is it an expression of familiarity, as has been claimed in the tradition perhaps beginning with Christopherson (1939)? Russell's (1905) analysis has prompted a considerable amount of debate, from Strawson's (1950) criticisms, to a more recent literature including Neale (1990), who argued for a quantificational analysis of the definite article, to Elbourne (2010), who argues against it, and who aligns himself with Heim's (1983) familiarity approach.

Another problem is to do with how we conceptualise the semantics. Russell works with quantification into the world, but cognitive semantics does not. That is to say that (crudely speaking) in a formal semantics, the meaning of a sentence is the fragment of reality it corresponds to, whereas within a cognitive theory, the meaning is a conceptual representation, anchored to reality by perception and memory. Given the theme of this volume, I want to ask how well a cognitive theory of semantics captures the differences and the debate in the formal theory, and whether, by taking a cognitive stance, we can sharpen up the question about how we should treat the semantics of the definite article.

In general, a cognitive theory of semantics helps with handling the semantic processes associated with grammaticalization: there is, after all, no reference to the

world; the whole of semantics is within conceptual structure, regardless of whether the bit of meaning that is under discussion is to do with sense, or reference, and so in a cognitive theory it is straightforward to account for changes that involve loss of content and more “referential” meanings. A cognitive theory should also make it possible to explain procedural meanings. This set of concerns is related to the issue of grammaticalization: we conceive of grammaticalization as a cognitive phenomenon, where semantic change is often intimately bound up with categorial change (Hopper and Traugott 2003). Furthermore, the cognitive theorist’s rejection of encapsulation makes it much more straightforward to track systemic changes within a cognitive account.

The final problem is to do with the heuristic value of the theory we adopt. How well does the theory help us find and understand new data? Generative theory gave us the tools to discover island effects, and the raising-control distinction. What similar heuristic value might a cognitive theory have? Within a cognitive theory, what kinds of semantic distinction come into focus because of the theory that has been adopted?

In this paper, I take the following line: I assume that the treatment of *THE* as a quantifier is right. There is evidence from a number of linguistic domains, but especially specificational sentences, that argues in its favour. However, there are some problems with modelling Russell’s specific proposals in a cognitive theory, because cognitive theories do not have unrestricted quantification into the world, so I argue that in a cognitive theory we need to develop a theory which is perhaps closer to the restricted

quantification of say Neale (1990), although a cognitive theory does not work with Neale's truth conditional semantics.

However, both Langacker (1991) and Davidse (2004) treat the definite article as a proportional quantifier (which Langacker calls a "relative" quantifier), and any theory needs a treatment of proportional quantifiers, so my tack is to argue that THE is a proportional quantifier, and to present an account of proportional quantifiers in Word Grammar.¹ I specifically formalise the account in Word Grammar because neither Langacker nor Davidse offer a formalisation, and I want to show how the analysis of THE as a proportional quantifier interacts with other parts of the grammar in as precise a way as I can, so as to make it possible for with scholars from different traditions to engage with the analysis.

Word Grammar is a cognitive theory of language which works with network representations. It was first reported in Hudson's (1984) monograph. The basic assumption of the theory is that language (and the rest of cognition, which language is embedded in) is a symbolic network. The theory is presented in Hudson (2007, 2010) and also in Gisborne (2010). I present such notions as the theory needs in the course of the paper when they become relevant.

So, to summarise, the questions I want to tackle are:

- What is the right semantics of THE to be able to capture the grammaticalization from the Old English demonstratives?

- How do we account for the right semantics within a (broadly speaking) cognitive theory of linguistic semantics?
- Are both treatments—familiarity and quantificational (or more specific formulations of them)—equally competent at accounting for the synchronic data?

There are several related points that I argue for in this paper which are listed here.

- I argue that a treatment of the definite article as a quantifier within a restricted theory of quantification is better than the familiarity account for handling the synchronic linguistic facts.
- I argue that restricted quantification can be theorized within a cognitive theory of language structure, and present some of Hudson's (2007) analyses.
- I extend these findings to THE, and explore some consequences of making these assumptions.
- Finally, I show that given some basic assumptions about the syntax, the quantifier approach makes for a simpler and more plausible diachronic story.

The paper has the following structure. The next section, §2, introduces the two main approaches to definiteness, drawing on Lyons (1999) and Abbott (2004); §3 discusses reference and the cognitive theory of reference; §4 explores the familiarity theory of definiteness in Word Grammar (the presentation being taken from Hudson 2007); §5

presents the restricted quantification version of Russell's theory of definiteness; §6 models the proportional quantifier approach in WG; §7 presents some linguistic reasons in favour of the quantifier approach; §8 discusses the two theories' merits with respect to grammaticalization; and §9 presents the results and conclusions.

2. Two approaches to definiteness.

There are two main approaches to definiteness (Lyons 1999; Abbott 2004). The familiarity theory is based on the idea that the referent is known to the addressee. It is generally attributed to Christopherson (1939), but is also adopted by scholars from as diverse theoretical persuasions as Bolinger (1977) and Heim (1983). The example in (1) gives a very simple example of how the familiarity account might be imagined to work.

(1) *A cat jumped into my garden ... the cat dug up my new bulbs.*

On the familiarity account, the indefinite phrase *a cat* establishes a referent, and so the familiarity of the cat is signalled in the second use of *cat* in the phrase *the cat*. By the second clause, *the cat* is familiar: we identify the bulb-digging cat as identical with the cat that jumps into my garden. The familiarity theory requires a psychological basis: Heim (1983) discusses the referential identity of the phrases *a cat* and *the cat* in terms of "discourse reference". Familiarity is located in the discourse, and not in the reference to a real cat in the world: I can say the sentence in (1) without knowing the cat or who owns it, and you can understand it, even though you have never been near my garden,

and do not know whether the cat is a marmalade tom or a tabby queen. Most linguists would agree that these discourse referents are intra-mental: if I say (1) to you, I create a cat referent in your mind, and reactivate it in the second clause.

The alternative approach is sometimes known as the uniqueness approach (Russell 1905). In Russell's account, a definite description (an NP with THE as the determiner) has properties in common with both the universal and the existential quantifier. The sentence in (2a) is analysed using the existential and universal quantifiers in (2b), and (2b) is given a prose translation in (2c).

- (2) a. *The student arrived.*
 b. $\exists x[Student(x) \ \& \ \forall y[Student(y) \ \forall \exists y = x] \ \& \ Arrived(x)]$
 c. *There is one and no more than one thing which is a student, and that thing arrived.*

Russell's theory has particular consequences. For example Russell claims that the sentence in (3a) asserts each of the propositions in (3b)-(3d). (Or, alternatively, asserts the conjunction of these clauses.)

- (3) a. *The King of France is bald.*
 b. *There is a King of France.*
 c. *There is only one King of France.*
 d. *This individual is bald.*

According to Russell, if any of the propositions in (3b)-(3d) is false, the whole sentence in (3a) is false. Thus, given that in 1905 there was no King of France, the proposition in (3b) is false, and therefore (3a) is false.

There is a challenge to Russell's approach to examples such as (3a). For example, Russell's analysis was attacked by Strawson (1950) who claimed (following Frege) that referring expressions **presuppose** a reference to something and so the existential clause in (3b) is simply a presupposition. In the terms of this theory, saying (3a) does not assert the existence of the King of France, it simply presupposes it. Therefore for Strawson an utterance of the sentence *the king of France is bald* does not have a truth value, because the definite description fails to refer. Strawson does not explain what a presupposition is, but there is a substantial linguistics literature from the 1970s that attempts to make sense of the notion of presupposition (Gazdar 1979, Kempson 1975, Wilson 1975), and at the very least we might concur that presupposition is likely to be a psychological phenomenon—loosely speaking we could say that a presupposition is something that speakers and hearers imagine or expect or believe to be the case about the world. Russell, on the other hand, is making an assertion about the relationship between the sentence itself, and the world.

Another challenge to Russell's approach comes from Hawkins (1978). Hawkins observed that there are several meanings of the definite article, and that the uniqueness approach does not really capture them. Hawkins's point can be easily confirmed by taking a look at a pedagogical grammar of English for the speakers of a

Slavic language. Such grammars explain not just the several meanings of definite articles, but also their many different contexts of use.

On the other hand, Russell's analysis was defended at length by Neale (1990), who explores the consequences for this theory in terms of the debate with Strawson, the behaviour of scope, substitutivity, and opacity, and the behaviour of definites as antecedents of anaphors. That is to say that where Strawson (1950) took a philosophical view of the debate, Neale (1990) located the debate in terms of how the theory addresses the analysis of particular linguistic data. A simple example can be seen in Milsark's (1977) treatment of definiteness effect phenomena, that is the difference between *there's a dog*² and *there's the dog*. There is a fundamental difference in interpretation between the two examples: the first sentence expresses the existence of the dog whereas the second is locative and identifies where the dog is. There are quite significant restrictions on the behaviour of definites in existential uses of THERE. What is more, there is a whole class of quantifiers that behave like THE, and which often cannot occur in the postverbal position of an existential sentence. These are the proportional quantifiers.

- (4) a. **There are most dogs.*
 b. **There are all dogs.*

If THE behaves like a proportional quantifier, then we need a theory that captures that observation, as well as the other generalisations that Neale (1990) identifies.

Now we run into two intimately related problems. The first is to do with the nature of quantification into the world, and the other is to do with the nature of reference. Russell's account, which was represented in (2b) exploits unrestricted quantification into the world: that is, the universal and existential quantifiers quantify over all objects in the world. As we will see in §5, we need a restricted theory of quantification to capture proportional quantifiers such as *MOST* which cannot be accounted for using unrestricted quantifiers. The other problem is that we cannot work with quantification into the world in a conceptualist theory, so how is a conceptualist theory to capture the analyses which rely on relations between sentences and the world?

The solution I shall adopt is to argue that cognitive linguistics needs a theory of linguistic reference, and that it needs to have a treatment of the proportional quantifiers such as *MOST*. The evidence in (4) suggests that it should be possible to extend the treatment of proportional quantifiers to *THE*, and so the main issue is simply what a cognitive theory of proportional quantifiers looks like. There is the question of whether the familiarity approach is the most obviously compatible with cognitive linguistics. Hudson (1990: 293-302, 2007: 226) assumes that it is, because familiarity is an obviously conceptual notion easily amenable to a cognitive analysis, in Hudson's case exploiting his conceptual binding relation. I argue here that the quantificational approach is equally amenable to a cognitive characterisation.

On the other hand, as I said in §1, Langacker (1991) and Davidse (2004) treat the definite article as a proportional quantifier (which Langacker calls a "relative"

quantifier), and the uniqueness criterion can apply to uses of *THE* which appear to be best handled in conceptual terms. If I say *the dog is growling*, and there isn't a dog immediately nearby, I must mean the dog that I live with. What's more, even if you and I are speaking on the phone, and you don't know that I live with a dog, you would conclude that I mean the dog I live with. The uniqueness criterion applies here (I share my home with 3 humans and only 1 dog)—and so arguably we require a conceptual notion of uniqueness.

I shall wrap up this section by pointing out that there are several ways of making an NP definite. Abbott (2004) gives the following list.

- | | |
|----------------------------------|---|
| • PRO | <i>Mary tried PRO to fly</i> |
| • Pronouns | <i>I, you, she, them</i> |
| • Demonstratives | <i>This, that, this chair over here</i> |
| • Definite descriptions with THE | <i>The king of France, the table</i> |
| • Possessive NPs | <i>my best friend's wedding, our house</i> |
| • Proper names | <i>Julia, Julia Child</i> |
| • Universally quantified NPs | <i>Each problem, every apple, all (the) girls</i> |
| • Generic NPs | <i>Pencils are made of wood</i> |

Abbott points out that this list, as she's presented it, is untheorized. The words which make NPs definite are not obviously all in a single class. For example, definite articles are known to grammaticalize out of demonstratives which are also a source,

diachronically, for relative pronouns. And it is not obvious that all of these different ways of expressing definiteness “mean” the same thing—given that THE has several meanings or uses, and given that there are several ways of expressing definiteness in Abbott’s list, we need to bear in mind that a complete story of THE and definiteness will include a wider range of meanings than I tackle in this paper.

In the next section, I discuss a cognitive theory of reference in order to get to an appropriate account of quantification in a cognitive theory.

3. Reference

Fauconnier (1994), Lakoff (1987), Jackendoff (2002), Hudson (1990) among others give plausible reasons why linguistic reference isn’t “in the world” or “in the fragment of the world that the sentence is describing”. The arguments I present here will be familiar to most readers, but it is worth revisiting them, because they lead to a particular stance on quantification.

In general, the view that semantics has to be treated as a domain of conceptual structure is common to several linguist theories, because one of the desiderata is a cognitively plausible computational system, and meanings (however they might be understood) have the property of being able to combine. We can think of this in terms of the distribution of NPs. A common first-year undergraduate syntax generalization is that English nouns cannot occur on their own, so (4a) is ungrammatical. The next example, though, is fine.

- (5) a. **Dog crossed the road.*
- b. *The/some/this/that dog crossed the road.*
- c. **The/a crossed the road.*
- d. *Some, this, that crossed the road.*

The generalization is that English common nouns, like DOG, need to occur with a determiner to make them grammatical. The generalization leads to a lot of spilled ink about whether the determiner or the noun is the head of the phrase, because most determiners, apart from the articles, can occur on their own as in (4d), with the ungrammaticality shown in (4c). One way to conceive of these facts is to say that the distribution of NPs is semantically determined: nouns have to have their definiteness value established in order to occur in an argument position. Nouns which are in a phrase with a determiner, or which are generic, have a definiteness value established.

This generalization has the virtue of not making the claim that all nouns need to occur with a determiner to be grammatical, so it includes the distribution of bare plural generics, and the distribution of bare predicative nouns as in . *We made her president.* However, crucially, this account of the behaviour of nouns and determiners and the distribution of noun phrases relies on an intramental notion of reference. Not only is definiteness a semantic rather than a syntactic property, but it is also a property of the semantics of reference rather than the semantics of sense. If we are going to capture the relevant facts to do with the distribution of NPs, we are going to need to include a semantic level of representation which is about reference within the grammar.

This is not surprising. I have already mentioned Heim (1983) and her use of discourse referents. Essentially, the argument is that the treatment of anaphora requires us to treat reference as a kind of concept, not as a relation to entities in the world. Take the example in (6).

(6) *Jane hates cats. Whenever she saw a cat_i in her garden, she shot it_i.*

The phrase *a cat* in (6) does not refer to any entity in the world. However, it is the antecedent to the pronoun *it*, which has to collect its referent from another word in the discourse. Heim (1983) argues, following Karttunen (1976) that we need a new notion of discourse referent for examples like (6), where *a cat* fails to refer (and so too must *it*, therefore). Similar arguments, about the need for representations in semantics in order to capture the relevant linguistic generalizations are made by Cann *et al* (2005).

But there are other reasons for adopting a mentalist view of reference. One is to do with identification questions. Jackendoff (2002: 301-2) observes that words such as *Manitoba* have indeterminate reference: is it a geographical space in Canada, a political region in Canada or a region on a map of Canada? Likewise a phrase such as *Beethoven's 9th symphony*—what can it refer to? The score? A particular performance? Obviously neither: imagine someone said the sentence in (7).

(7) *Beethoven's 9th symphony is predictable and boring, but I still get a thrill when the choir blasts out "Ode to joy".*

You cannot argue that the phrase *Beethoven's 9th symphony* does not refer. As the subject of the clause it must. But what is being referred to is not anything in the world—it's something in the speaker's head (and the hearer's), their concept of the symphony, a generalization over a number of experiences of it (which might include listening to it live, to recorded versions, to performing in it as a singer or as an instrumentalist or as a conductor, to reading the score).

There are, then, two arguments that “referents”—or at least the referents which are relevant to linguistic, and especially grammatical, description—are concepts. There is a third argument for abandoning a realist view of linguistic meaning in favour of a conceptualist one, and that is the argument from classification. Lakoff (1987) argues that human cognition classifies according to prototypes, and that Aristotelian categories, which are organised around set membership, fail to capture these differences. Categorisation is relevant, because if word meanings aren't sets (because categories aren't hard and fast) then we need to decide how categorisation is done, and how it interacts with other aspects of meaning. Most cognitive theories assume default inheritance—which is a kind of logic originally designed to capture the non-monotonic properties of human reasoning. (Langacker calls default inheritance “schematicity”.) Inheritance gives us a way of capturing both prototype effects and how word meanings combine to form complex concepts.

These observations about reference mean that there is no way in which external quantification, analysed with sentence connectives, in a first-order predicate logic can

make any sense in a cognitive theory. In a cognitive theory, the relationship between the representation and the world is mediated by the apparatus of cognitive psychology and perception. Examples such as (8) are not translatable into a cognitive theory in a straightforward way, because (8) quantifies over all things in the world: it means for all things in the world x , if x is a person then x dies. A cognitive theory has no mechanism for capturing quantification over a variable with a range which is every single entity in the world.

$$(8) \quad \forall x \text{ Person}(x) \rightarrow \text{die}(x)$$

So how do we capture the kinds of information that are described using external quantification such as (8)? It depends, of course, on the theory, and on how it models linguistic representations. I will give an answer using Word Grammar: the main point being that all of the representations in WG are conceptual structures. The theory claims that language is a symbolic network. Naturally enough, WG does not use external quantification (i.e. there is no quantification over variables) because all of WG's semantics is intra-mental, and not understood in terms of a relationship between the utterance and the fragment of world it describes. Quantification over variables is intended to establish which object in the world is under discussion. Hudson (2007: 31) gives the example in Figure 1 as a WG representation of (8).

Figure 1 about here

Figure 1 presents its own complexities. It states that the category 'Person' is the agent of dying of an instance of 'die'. WG states its equivalent of universally quantified statements using default inheritance: anything that inherits from 'Person' automatically inherits the property that they die; however anything that inherits from 'die' will not inherit being a person. If we wanted to state that all, and only, animals and plants die, we would need to have a link from the 'die' node at the top of the diagram.

There are two kinds of relation in a WG diagram. The line from 'die' to '1' with an upside down triangle at the top represents the Isa or "is an instance of" relation, which is the relationship of default inheritance. Therefore, this part of the diagram says that '1' is an instance of 'die'. The second kind of relation is shown by an arrow, which can be thought of as a function from an argument ('1' in this case) to a value ('Person').

4. The familiarity theory of definites

In this section, I want to offer a particular account of the familiarity theory, which is generally assumed by functionalists and cognitive linguists (although not Langacker 1991, Hawkins 1978); note though that there are formal approaches which also assume a familiarity view, including Dynamic Syntax (Cann *et al* 2005) and File-change Semantics (Heim 1983).

This approach assumes that a definite NP is "Hearer Old". It is straightforward to show how this approach to definiteness works out in a formalised cognitive theory,

because Hudson (1990, 2007) assumes this approach for Word Grammar. In his most recent formulation, Hudson (2007: 226–227) assumes “conceptual binding” which is a way of formalising familiarity in a conceptual network. Conceptual binding works both for definite NPs (which are bound to a concept opened up in the conceptual network by earlier discourse) and reflexive pronouns (which have particular constraints on their distribution). Essentially, the claim is that the semantics for THE inherits a particular link which has to connect to an antecedent concept. If there isn’t an antecedent concept for this link to attach to, the utterance is ill-formed.

Hudson’s bound concepts are related to Prince’s “Hearer Old information”, which Abbott (2004) describes in this way:

The concept of familiarity which Christophersen has in mind here seems quite similar to Prince (1992)’s concept of HEARER-OLD INFORMATION, which she aligns with the idea of information which is ‘in the permanent registry’ (Kuno 1972), or ‘culturally copresent’ (Clark and Marshall 1981).”

Hearer-old information subsumes discourse old information. Hudson’s analysis of THE is given in Figure 2 below.

Figure 2 about here

The best way to explain the diagram is to quote from Hudson’s own account.

Definiteness comes from the determiners, which relate directly to the noun’s referent.

In 1990 I interpreted definiteness in terms of the addressee’s knowledge (Hudson 1990:

293-302), with definite referents already known to the addressee and indefinites new to the addressee. This may be an accurate analysis, but a much neater way to explain the difference is now available thanks to the binding mechanism ... definite referents are bound. This is not only much easier to show in the analysis, but it also explains why the addressee is expected to know the referent already. This binding is shown by the double arrow which means 'directed identity' linking one obligatory node to another, and is only 'potential' in this diagram because the object of the binding (the antecedent) is not available in this sentence. It must be available in the interlocutors' minds, so, at least in principle, we could complete the binding if we had a network analysis of the relevant parts of either the speaker's or the hearer's mind. (Hudson 2007: 226)

Hudson's account is explicitly cognitive—he works with a notion of what is in the speaker's and hearer's minds, and how the relevant parts of their conceptual structures are activated in the course of a discourse. This description of a familiarity approach shows how it links to a cognitive theory, and also presents a formally precise account. We could look at other formal accounts—Heim (1983) seeks to rehabilitate the familiarity theory within File Change Semantics as does Elbourne (2010), and Kamp (1981) presents an alternative Discourse Representation Theory account—but as these theories are not cognitive, they are not directly relevant to my concerns here. I argue in the section below that the quantificational theory of definiteness also accounts for the notion that a definite NP is familiar.

5. An alternative theory of definites

Neale (1990) presents a set-theoretic treatment of the definite article, located in the Generalized Quantifier Theory which accounts for proportional quantifiers such as *MANY* and *MOST*.³ It has been widely noted that unrestricted quantification analysed with the standard sentence connectives cannot help with words such as *MANY* or *MOST*, which therefore have to be analysed in an alternative system. Neale extends that system to the definite article. Neale's analysis (1990: 40-43) follows a treatment in Wiggins (1980) and Barwise and Cooper (1981). Neale's argument is given here (1990: 40).

[A]s soon as we encounter a genuine binary structure, we get stuck. For instance, suppose we wish to represent (7):

(7) Most men are immortal.

What we require is a formula of the form of (8).

(8) (MOST x) (man x © immortal x)

where '©' is a binary, truth-functional connective. Clearly '©' cannot be '&' for then (7) would mean that most things are men-and-immortal. Nor can '©' be '⊃' for then it would mean that most things are if-men-then-immortal. But since nearly everything is not a man, nearly everything is if-men-then-immortal; therefore the sentence will come out true whether or not most men are immortal (here I borrow heavily from Wiggins' succinct discussion). In fact, there is no sentential connective that captures what we

require of ‘@’; indeed, it is not possible to define ‘most *Fs*’ in first order logic at all, even if attention is restricted to finite domains.

The problem is that in (8) the “quantifier” *most* is ranging over the entire domain of quantification rather than just those things that are men. Intuitively, we want something like the following result:

‘most *Fs* are *Gs*’ is true if and only if $|F \cap G| > |F - G|$

(F = the set of things that are *F*. $F \cap G$ = the set of things that are both *F* and *G*. $F - G$ = the set of things that are *F* and not-*G*. $|F \cap G|$ is the cardinality of $F \cap G$. This means that we should treat ‘most’ and other natural language determiners (‘some’, ‘every’, ‘all’, ‘no’ and so on) as exactly what they appear to be: devices that combine with two simple or complex formulae (or predicates, depending on how one views matters) to form a formula.

To put it another way, *MOST* should be treated as a **restricted quantifier** rather than as a fully general quantifier. For *most men are mortal* what we need is for the whole set of men to be sorted or given at the outset. This way *most* will pick out a proportion from the set of men, rather than from the set of everything or objects in general. Rather than creating a complex proposition using connectives, the range of the quantifier *most* is restricted by the noun *men*.⁴

(9) *Most men are mortal*

[Most x : men x] (mortal x)

Neale (1990: 42-3) shows that all quantifier determiners can be presented in the same way. I give some of his truth clauses immediately below. After presenting his truth clauses, Neale goes on to show how restricted quantification can capture quantifier scope phenomena as well as unrestricted quantification, and how the Theory of Descriptions can be presented in the language of restricted quantification.

Truth clauses for EVERY, NO, SOME, AN and MOST. (From Neale 1990: 42-3)

(*1) '[every x : Fx] (Gx)' is true iff $|\mathbf{F} - \mathbf{G}| = 0$

(*2) '[no x : Fx] (Gx)' is true iff $|\mathbf{F} \cap \mathbf{G}| = 0$

(*3) '[some x : Fx] (Gx)' is true iff $|\mathbf{F} \cap \mathbf{G}| \geq 1$

(*4) '[an x : Fx] (Gx)' is true iff $|\mathbf{F} \cap \mathbf{G}| \geq 1$

(*5) '[most x : Fx] (Gx)' is true iff $|\mathbf{F} \cap \mathbf{G}| \geq |\mathbf{F} - \mathbf{G}|$

Truth clause (*1) can be translated into natural language prose as, "every x which is an F is also a G if the cardinality of F minus G is zero." That is, if you like, the material on the

left of the predicate “is true” is the syntax of the quantification, and the material on the right is the semantics, in a theory where the semantics consists of the conditions under which a statement is true. To put it another way, if you assume a restricted quantification, “The semantics of definite descriptions can then be given in terms of generalized quantifier theory” (Ludlow 2007), which is to treat them as proportions of a set. Generalized quantifier theory claims that natural language quantifiers express relations between sets. This can be shown diagrammatically, as in figure 3, below.

Figure 3 about here

Neale goes on to argue that the restricted quantifier treatment allows a straightforward handling of quantifier scope interactions, which obviously any theory needs to address, and then presents a version of Russell’s Theory of Descriptions, pointing out that the Theory of Descriptions does not require Frege’s unrestricted quantification, and that to present it in a restricted formalisation is simply to “choose a language other than that of *Principia Mathematica* in which to state and apply the theory.” I will come back to the representation of THE in §6; before I get there, I want to address the issue of the semantics of a restricted quantification in a conceptualist theory.

In cognitive theories, we don’t treat either verbs’ meanings or nouns’ meanings as sets. Furthermore, in a cognitive framework, the conditions under which a sentence is true do not make up its meaning. The assumption is that the relationship between a proposition and the world is mediated by perceptual, embodied experience, and that

the meanings of sentences are conceptual representations. Therefore, we can borrow from Neale the notion of working with a restricted set, but we have no need of any statement of the conditions under which a proposition is true. Therefore, in a cognitive theory of quantification, we need to discuss quantifiers such as *MOST* in terms of how it scopes over a set. This is part of a mental model. The plural noun denotes a set, and *MOST* establishes a scoping relation over that set, therefore expressing a proportion: a subset of the set denoted by the plural noun. Following Neale, we can treat *THE* as a proportional quantifier, as well as the quantifiers given above in (*1) to (*5).

Recall that in §2, I pointed out that Milsark shows that so called definiteness effects in existential *THERE* sentences are actually effects that single out proportional quantifiers, so it makes sense to treat *THE* as another proportional quantifier. The claim in (2b), repeated here as (10), is that a definite description denotes a set, which is understood both existentially and exhaustively.

$$(10) \quad \exists x[Student(x) \ \& \ \forall y[Student(y) \ \forall \exists y = x] \ \& \ Arrived(x)]$$

The formula in (10) asserts the existence of the content of the set, it quantifies over all members of the set, and therefore it argues that the existence claim of Russell's *the king of France is bald* is actually an assertion, not a presupposition. Essentially, the point here is to do with whether definite expressions have a truth value or not. For Russell, *the king of France is bald* is false; for Strawson, it is neither true nor false. If the existence of the King of France is presupposed, then the existence claim is not part of the semantics

(although see Burton-Roberts 1989 for an interesting alternative position on presupposition). In §7.3, the argument that the existence claim is part of the semantics is part of the analysis of specificational sentences that I sketch.

Neale (1990: 45, 46) gives (11a) for singular definites with THE and (11b) for plurals.

- (11) a. '[the x : Fx] (Gx)' is true iff $|F - G| = 0$ and $|F| = 1$
 b. '[the x : Fx] (Gx)' is true iff $|F - G| = 0$ and $|F| > 1$

Neale points out that '[the x : Fx] (Gx)' is *definitionally equivalent* to $(\exists x)((\forall y) (Fy \equiv y=x) \text{ and } Gx)$. He also points out, following Chomsky (1975), that the relationship between singular and plural definite descriptions comes into view on the Generalized Quantifier Treatment given in (11), because the difference is simply one of cardinality.

I haven't yet explained how treating a definite expression as containing a proportional quantifier captures the familiarity effect. We can understand this by thinking about incomplete descriptions. Kearns (2000: 97) gives the example in (12a), which has the meaning given in (12b).

- (12) a. *All men must report before taking leave.*
 b. *'All enlisted men now serving on this base must report before taking leave.'*

The example in (12a) is incomplete out of context, whereas (12b) can be used out of context because the description is complete. But (12a) is the normal expression because “the speaker or writer assumes, or presupposes, that the audience can identify the background set, either from general shared knowledge, or because the information has been given earlier in the discourse” (Kearns 2000: 80). As Kearns says (2000: 96), “To understand a proposition with a strong quantifier, the hearer must be able to identify the background set.”⁵ To put it another way, familiarity falls out of the quantificational story because an NP with a strong or proportional quantifier expresses a quantification over a background set, identifying the relevant subset—in the case of a singular definite NP, the relevant subset is a singleton set.⁶

My claim in the next section is that we can develop a conceptualist or cognitive version of the restricted set treatment of proportional quantifiers (including *THE*) and as a result compare the familiarity and the quantifier treatment of definite articles from within the terms of a single framework. The set theoretic treatment of *THE* is predictive: if it is right, it should be possible to find linguistic contexts where a definite description has to be understood as denoting a set rather than as having a referent. I set out to show such a situation in §7. In the next section, I show how a version of this theory can be modelled in a cognitive network.

6. Modelling the quantifier theory in a cognitive theory of language structure

Any cognitive theory needs to be able to model the meanings of words such as ALL, EACH, EVERY, SOME, MOST, MANY. If the representations in (11) are reasonable ways of modelling THE, then it should be possible to model THE in a way analogous to a restricted set in any cognitive theory that has a way of modelling a proportional quantifier like MOST. Hudson (2007) gives representations for cardinal sets (*two dogs*) and for EVERY. EVERY is a bit complicated, because it's singular in the morphosyntax, and semantically plural, so I shall not present those analyses here; instead, I shall work up an analysis of proportional quantification by looking at the semantics of conjoined NPs. Hudson (2007: 34-35) presents analyses of the three sentences/interpretations in (13).

- (13) a. *John and Mary bought a house together.* (collective)
 b. *John and Mary each bought a house.* (distributive)
 c. *John or Mary bought a house.* (arbitrary member of the set)

The diagrams that follow are taken from Hudson (2007: 34–35); the first diagram shows an analysis of the collective interpretation of *John and Mary bought a house*. I take each of the diagrams in turn, and explain how they capture the intended meaning.

Figure 4 about here

In the diagram, the agent of the 'buying' event (the arrow labelled "Er") has as its argument a concept which is a set.⁷ The set is defined explicitly: its nature as a set is

shown by the inheritance link from the category 'set' to the '1', which indicates that there is just one set that is the Er of the 'buying', showing that there is only one act of buying. The set is also defined ostensively, in that it is the instance of the type 'set' which has 'John' and 'Mary' as its members and which is the Er of this particular instance of the type 'buying'. In the latter way, the diagram shows that the buying event is a token of 'buying', not the type itself. In summary, the diagram says that the house was bought by the set of John and Mary.

A distributive interpretation is a bit more difficult to represent. In order to capture distributive conjunction, we need to show that it is both members of the set that buy the house—and that there are two buying events, therefore.

Figure 5 about here

The diagram says that the Er of 'buying' is the typical member of the set; the set is, like the previous set, defined as a set with two members, John and Mary. Since both John and Mary are instances of the typical member (John *Isa* typical member and Mary *Isa* typical member) they both inherit the relation to the Er of the event. The diagram does not permit me to show the way in which there are two events; in order to do this, it is necessary to show more of the verb's semantics as well. In Figure 6, I add this next part of the complexity. Figure 6 draws on the analysis of *Two researchers wrote three articles* in Hudson (2007: 230).

Figure 6 about here

In Figure 6, the referent of *bought* is shown as the typical member (i.e. any instance) of a set which *Isa* (is an instance of) a buying event where John *Isa* the typical member and Mary *Isa* the typical member. In this case, the cardinality of the set is defined by the cardinality of the set of house buyers so, just as there are two buyers, there are two buying events. The analysis relies on WG's intramental notion of reference, which allows verbs to refer, so that the conceptual structure permits the same tools to quantify over events and things.

The previous two diagrams force a universal quantification, because both John and Mary are involved in the house buying, either jointly or separately. What if we have a set conjoined by OR rather than AND? Hudson (2007: 35) says, "The effect of changing *and* to *or* is much the same as that of changing universal to existential quantification, because we change from 'every member' to 'some member'." As we shall see, this difference is important because *THE* forces both a universal and an existential quantification, whereas *A(N)* only involves existential quantification.

Figure 7 about here

In Figure 7, an arbitrary member of the set is chosen as the agent of 'buying'; I have shown that arbitrary member with a dotted line to show that it is not the same as the other two 'member' relations: 'member 1' and 'member2' define the set—the set is the

set of John and Mary. The third member relation, ‘m3’ in the diagram is just an arbitrary member that is bound to one or other of the actual members, but crucially neither of the actual members is a m3. The dotted line is a notational convenience to show this arbitrary property.

The way the existential quantification works is through inheritance: “any node X always means ‘every X’, [...] if X is a Y, then it is merely ‘some Y’, so its properties are not inherited by other instances of Y” (Hudson 2007: 33). By linking the Er of ‘buying’ to an instance of a relation which is a member of the set, the diagram shows that the buyer is one of John and Mary—as long as the “1” which is the Er of the ‘buying’ event, and “m3”, is bound to John or Mary.

So how do we extend this to THE? One way is to exploit the treatment of sets developed for collective and distributive conjunction above, and for OR, and to build a semantics in set terms. This would give a WG account of Neale’s restricted quantifier treatment of definite articles, which could be extended to other proportional quantifiers. Although our intramental semantics does not allow us to exploit Neale’s truth statements, we can, however, also take an insight from them, and set a cardinality statement on our sets for singular and plural THE: in this way, like Neale I should be able to use the same basic representation for singular and plural THE, with the difference between them being limited to a statement of set size.

However, there is a small wrinkle: like conjoined NPs, plural NPs determined by THE can have either collective or distributive interpretations. If I say *the researchers wrote a paper*, you can interpret it either as (a) they wrote the paper together, or (b)

they each wrote a paper, so in the end there were two papers. We need, therefore, to propose a semantics which captures this difference. This will be an advance on Neale (1990), which does not show the difference between collective and distributive plural definite NPs.

The first representation, in Figure 8, gives a set representation of THE; I analyse the sentence *The cat played*. My diagram says that the Er of 'playing' is a set, which has one cat as its member.

Figure 8 about here

For the diagram to capture the insights of Russell's (1905) analysis, it has to capture both the universal and existential commitments which are expressed in (3), the argument being that the sentence *the King of France is bald* asserts the existence of the King of France and that there is only one King of France. The universal quantification is achieved by stating that it is the set which is the Er of playing: if it is the set, then all members of the set are quantified over. The set is defined ostensively as the set which has one member, which is the cat. Existential quantification is achieved by the Isa relationship between the concept 'cat' and the node which is the member of the set. The diagram therefore gives us a network analysis which captures the same insights as Neale's restricted quantifier approach.

In the next diagram, I tackle the collective interpretation of *the cats played*—an interpretation that says that the cats played together. This interpretation asserts that there is only one playing event.

Figure 9 about here

Like the previous diagram, this analysis shows the Er of ‘playing’ as the set of cats. The diagram is different from the diagram for a singular definite NP in that the set-size is shown to be greater than one. However, it captures the universal and the existential quantification in the same way as the previous diagram. The difference between a singular definite NP and a plural definite NP with collective interpretation is established by the cardinality of the set. The next diagram presents a distributive analysis of plural definite NPs.

Figure 10 about here

The diagram works in the same way as the diagram for the distributed interpretation of *John and Mary bought a house (each)* in Figure 6: the set has a typical member which links to the Er; the actual members of the set all Isa this member. In this diagram, I have not shown the actual members—the fact that this is a set with more than one member is shown in the ‘size’ relation. The event is also given as a set to show that what is

intended is a distributive interpretation where there is a separate playing event for each cat.

So far, I have shown how it is possible to have a set-based account in Word Grammar—a cognitive linguistic theory—of the definite article in its three main interpretations: within a singular NP, within a plural NP interpreted collectively, and within a plural NP interpreted distributively. What about the indefinite article $A(N)$, and its plural counterpart $SOME$?

It is obvious that $A(N)$ needs a treatment as a quantifier. It interacts with other quantifiers in the scope phenomenon involved in the two interpretations of the sentence in (14a) given in (14b) and (14c) from Neale (1990: 119).

- (14) a. *Every man danced with a woman who was wearing Ray-Bans.*
b. *[every x: man x] ([a y: woman y & y was wearing Ray-Bans] (x danced with y))*
c. *[a y: woman y & y was wearing Ray-Bans] ([every x: man x] (x danced with y))*

On the interpretation in (14b), there are several different women wearing Ray-Bans, and the men are dancing with different women; on the interpretation where *a* has wide scope in (14c), there is just one woman wearing Ray-Bans, and every man dances with her. $A(N)$ must be a quantifier: it behaves just like one with respect to scope alternations.

SOME is straightforward: it occurs with plural nouns (in several dialects of English, SOME, like THE, can occur with both plural and singular nouns, but I shall only take the plural case here). I am not going to propose diagrams here, because they would become unfeasibly complex, and because my main focus is on showing how we could capture definiteness set-theoretically in a cognitive theory, but we can at least think about how we could analyse SOME.

It seems that the main point is that SOME involves a relation between sets. When SOME has a full vowel, we can analyse *some cats* as analogous to *some of the cats*, in which case we need to identify the referent of *some* as a subset of the set of cats.⁸ Because the particular subset is not known (*some cats* is indefinite) it is an arbitrary subset of members of the set of cats that is identified by *some cats*. The treatment of SOME belongs in a larger scale treatment of quantifiers in WG, which is the subject of further research. However, the subset relation will figure in the treatment of all proportional quantifiers apart from those which, like THE, involve universal quantification.

7. Comparing the familiarity theory with the quantifier theory

In section 6, I showed how a cognitive theory of linguistics could capture the determiner-as-quantifier analysis. Here, I want to argue for the different merits of the analysis presented in §6, and the alternative familiarity analysis presented in §4. There are three case studies where it has been claimed that the quantifier treatment of

definite descriptions captures their grammatical behaviour better than the familiarity theory.

7.1 Case study 1: scope effects

In this case study, we look at how definite descriptions behave with respect to scope effects. We find that they behave like quantifiers—in brief that definite descriptions can have narrow scope interpretations as well as wide scope interpretations. The main arguments are presented in Neale (1990: 118-164), although there have been rejoinders to Neale's claims, such as Elbourne (2010). Neale's arguments are long, detailed and engaged with a long history so here I only sketch the lineaments of a few of his points.

First, Neale points out that in a simple way definite descriptions do not behave like $A(N)$ with respect to scope alternations. A sentence such as (15) will always be interpreted as if *the woman who was wearing Ray-Bans* took wide interpretive scope (i.e. as if it identified a single individual with whom each man dances).

(15) *Every man danced with the woman who was wearing Ray-Bans.*

In this respect, THE behaves quite differently from $A(N)$. However, Neale shows that THE does, in fact, enter into scope interactions which argue in favour of its being a quantifier, as in *each man danced with the woman who was sitting next to him*. For example, in a sentence such as (16a), the default interpretation is that *each girl* has a

different mother from the other girls in the set; likewise, in (16b), every man has his own wife. In these cases, the definite expression cannot take wide scope.

(16) a. *The mother of each girl waved to her.*

[each *y*: girl *y*] ([the *x*: *x* mother-of *y*] (*x* waved to *y*))

I.e. there is a different mother for each girl.

b. Every man *respects the woman* he *marries*.

[every *x* man *x*] ([the *y*: woman *y* & *x* marries *y*] (*x* respects *y*))

I.e. there is a different wife for every man.

Neale argues through a range of example types and different contexts, including modal contexts; I will not take all of them, but I will present one further argument from the interaction of definite descriptions and modal contexts (Neale 1990: 121).

(17) *The first person in space might have been Alan Shepard.*

This example has the two interpretations in (18) and (19).

(18) *possibly [the *x*: first-person-into-space *x*] (*x* = Alan Shepard)*

(19) *[the *x*: first-person-into-space *x*] possibly (*x* = Alan Shepard)*

The interpretation in (18) is true: there is some counterfactual state of affairs in which Alan Shepard was the first person into space; however the other interpretation is false because it asserts that there is some counterfactual state of affairs in which Yuri Gagarin is in fact Alan Shepard.

We can set truth conditions aside, or reduce them to a handy way of establishing that the two interpretations actually mean different things. Irrespective of how we assign meaning, we can agree that scope interactions are a linguistic phenomenon, and that therefore they need to be accounted for—including in a cognitive theory. Once we agree that, we are in the business of arguing about THE as a proportional quantifier—like other proportional quantifiers, given an appropriate context it interacts in scopal phenomena. It is incumbent on the person who asserts that THE is not a quantifier to explain these scope phenomena by some other means.

7.2 Case study 2: the definiteness effect

I mentioned this argument in §2; it follows arguments from Milsark (1977).⁹ It follows from a discussion of what can be in the complement position of a THERE BE existential construction. Definite descriptions behave like “strong NPs” with respect to THERE BE contexts. This is known as the definiteness effect. NPs with proportional quantifiers such as ALL and MOST are strong.

- (20) a. **There were all cows in the field.*
b. **There was the cow in the field.*

Weak NPs, on the other hand, such as those quantified by a numeral or A(N) are permitted in the existential construction, so the examples in (21) are fine.

- (21) a. *There was a cow in the field.*
b. *There were three cows in the field.*

The data generalization captured by the difference between (20) and (21) argues in favour of classing THE among the proportional quantifiers, especially when it is taken together with the data in §7.1, and the material presented next in §7.3.

7.3 Case study 3: specificational sentences

One area of research which has been particularly intractable has been the behaviour of specificational sentences, such as the examples in (22). The examples are from Patten (2010).

- (22) a. *The murderer was John.*
b. *The one that murdered Sally was John.*

These sentences have a particularly complicated set of properties: they involve focus—the postverbal name being focused here—and they appear to involve non-standard predication, where the NP immediately before the copula is a predicate and the NP after

is its argument. Variants of the latter claim have been made in the literature, for example in Moro (1997) and Mikkelsen (2005). Other scholars, such as Heycock and Kroch (1999) have argued that specification sentences are a kind of equative construction where both NPs refer, and are identified as co-referential, and the focus effects are established separately from the equative analysis.

Both approaches have their downsides. The inverse predication approach of Moro and Mikkelsen involves a considerable amount of syntactic machinery bringing about movement. The equative approach of Heycock and Kroch requires postverbal movement in the semantic representation.

The alternative approach, which Patten (2010) offers, involves no syntactic machinery, and no LF movement. Furthermore, it does not require any kind of level or representation of information structure. Patten makes two claims—that in the specificational sentence in (22a), the NP *the murderer* denotes a set, and the post verbal NP *John* is the extension of that set. The analysis is only possible on an account where THE is a quantifier; without this treatment of definite descriptions, it is not possible to have the linguistic analysis. Because definite descriptions express both existential quantification and universal quantification (over a restricted set) the analysis gets both the existential commitment of specificational sentences and their exhaustiveness.¹⁰

This last fact explains the data in (23).

- (23) a. *The murderer was John.*
b. *The murderers were John, Fred, and Alfie.*

c. **All murderers were John, Fred, and Alfie.*

d. **All the murderers were John, Fred, and Alfie.*

The question posed by the examples in (23) is this: if *THE* involves universal quantification over a set, why are the examples in (23c,d) ungrammatical, whereas the examples in (23a,b) are not?

The answer is simple: what makes the specificational sentence specificational is not just the universal quantification over the restricted set, but is also the existential commitment. *ALL* does not involve existential commitment, and so (23c) is ungrammatical. In the case of (23d), the story is more complex because of the *the* under *all*. Here, we can invoke the earlier discussion of proportional quantifiers in §6 above, where I pointed out that they necessarily involve relations between sets. The phrase *the murderers* in (23d) is a subset of *all the murderers*, and it is not possible for the extension of a subset of a set to be identified in such a construction: the extension of the *all* set is the set *the murderers*, not the members of the set of murderers, and the existence of that set is not asserted by *all*.

Patten extends this account to cleft clauses, such as those in (24).

(24) a. *It was John that was the murderer.*

b. *It was Cicero who once said, "Laws are silent at times of war".*

The argument is that in (24a), for example, *it...that was the murderer* is a discontinuous definite NP, and *John* is the extension of the set that the discontinuous NP describes. The account extends nicely to the grammatical emergence, or constructionalization of the cleft construction, as described in Gisborne and Patten (2011).

As well as grammaticalization, this analysis of cleft clauses has another very simple benefit. Patten (2010) presents her analysis in construction grammar, and her account is both simple (there are no derivations) and compatible with cognitive assumptions, once it is agreed that a treatment of THE as a proportional quantifier makes sense in a cognitive theory. Patten does not present a syntax of the specificational sentence, but we can assume that the lefthand NP is simply a syntactic subject (it can raise over SEEM as in [25a]) and that the righthand NP is simply a complement.

- (25) a. *The murderer seemed to be John.*
b. *Which murderer did she think was John?*
c. **Which murderer did she think that was John?*

Note, in fact, that the left hand NP not only raises over SEEM but is also subject to the THAT-trace effect, where a subject in English cannot be extracted over THAT.

There is no need on this analysis to posit leftward movement of the predicate NP—the apparent predication falls out of the semantics. There is no need to posit a level of information structure which captures the universal quantification and existential

commitment because they are in the semantics anyway. We can capture the syntactic properties of specificational sentences and their conflict with the semantics as a simple constructional mismatch.

Patten does not give a formal account of specificational sentences, which in a formal semantics would require type-raising of the definite NP to a predicate in order to get compositionality. I present a WG analysis of the structure of a specificational sentence in the next diagram.

Figure 11 about here

In the diagram, the singular NP has the same structure as I provided above in Figure 8. The set treatment of *THE* gives us the relevant semantics, therefore providing the existential commitment and exhaustivity. The syntactic structure is one of complementation, not predicative complementation, and in a complementation structure with *BE*, you expect an identity relation—as in *Tully is Cicero*. However, it does not make sense to identify a set with an individual, although you can identify a member of a set with an individual. Therefore, the semantics that works is one where the member of the set is identified with the referent of the second NP.

Essentially, the analysis differs from an equative analysis only inasmuch as the first NP is given a set-based treatment as a definite NP. The rest of the structure is what we expect for *BE* with an NP complement, which needs there to be an identity relation in examples such as this.

I finish this section by noting that the familiarity analysis cannot explain the quantifier scope interactions or the definiteness effects, and that the analysis of specificational sentences I have just offered is only possible in an account where we treat THE as a proportional quantifier.

8. The theories and grammaticalization

In this section, I argue that a set-theoretic treatment of definiteness, and therefore of THE, is compatible with the grammaticalization facts and makes for a simpler story than the alternative familiarity approach. In making this argument, I rely on a particular syntactic assumption, that in the “NP” it is, in fact, the determiner that is the head, and that definite articles are, like demonstratives, just transitive pronouns. The difference between a definite article and a demonstrative is that the definite article is obligatorily transitive, whereas the demonstrative is optionally so. This version of the determiner-as-head hypothesis is from Hudson (1990: 268-276); it differs from Abney’s (1987) version, which assumes a different position on the nature of the superordinate category, and which assumes a different structure within the NP. That is to say that I am not adopting Abney’s version of the determiner-as-head analysis, because Abney assumes that pronouns belong in the determiner category; I am assuming that determiners belong in the pronoun category. The Old English demonstrative SE is the source for modern English THE (Mitchell 1985, Traugott 1992), and this is often treated as a prototypical grammaticalization path (Traugott 1982, Lyons 1999).

The story I offer is a simple one: I take it that the semantics for THE that I have argued for, and worked out a WG/Cognitive representation for, is the most schematic of the meanings that definite expressions might have. If we recall Abbott's list of definite expressions in §2, it cannot be assumed to be the case that definite expressions only mean universal and existential quantification over a restricted set. Take 's: this is definite, and so must have the set-theoretic semantics of THE, but it also means possession. The identity of the possessor has to be separately given, of course, but nevertheless, possession is additional meaning that is not present in THE. The grammaticalization of THE does not involve categorial reanalysis; nor does its requirement to have a complement noun. The apparent categorial reanalysis and the complementation pattern are entirely explainable in terms of a very simple semantic change, so that essentially the semantic content of THE reduces to the quantifier analysis I have given.

We can start with Diessel's (1999: 129-129) presentation of the grammaticalization of the English definite article from the demonstratives. Diessel writes,

The use of anaphoric demonstratives is usually confined to non-topical antecedents that tend to be somewhat unexpected, contrastive or emphatic ... When anaphoric demonstratives develop into definite articles their use is gradually extended from non-topical antecedents to all kinds of referents in the preceding discourse. In the course of this development, demonstratives lose their deictic function and turn into formal markers of definiteness. An example of such a definite marker is the article *the* in English.

He adds that there are three factors involved in this process.

- “demonstratives lose their status as free nominal when they become reanalysed as definite markers”
- “demonstratives are significantly more often inflected than articles, which suggests that adnominal demonstratives often lose the ability to inflect when they grammaticalize as definite markers.”
- There is a cline: demonstrative>definite marker>specific indefinite marker

How do the two theories handle the grammaticalization facts?

Because demonstratives are deictic, we need to start with thinking about deixis.

The WG analysis of deixis relies on the situated nature of the sentence, or utterance:

because WG assumes that utterances are intramental, there is no distinction between utterance and sentence, both are conceptual tokens. We can therefore make the following claims about deixis.

- Words in utterances have a time, place and speaker.
- When I utter *me*, there is a relationship between me as speaker, the place and time of the utterance, and the discourse context.
- This is all modelled in conceptual structure.
- Deictic expressions are therefore subjective.

How does this play out in the semantics of the demonstratives? Let us take *THAT* as a case study. There are, I think, two main points: first, the demonstratives have to inherit the definite semantics of *THE*. This is not to say that *THAT* is *THE*, but to say that the meaning of *THAT* must inherit whatever it is that makes *THE* definite, because demonstratives are also definite. This observation suggests that the meaning of *THAT* is definite in the ways that the meaning of *THE* is definite, plus it has an additional layer of deictic meaning. The claim is very similar to what I claimed above for *'s*, that it inherited the definite semantics, with an additional layer of possessive meaning.

Figure 12 presents a WG analysis of *THAT*, showing both its definiteness and that it is deictic.

Figure 12 about here

The diagram in Figure 12 is slightly underspecified, in that I have not shown the difference between *THAT* when it is anaphoric and *THAT* when it has an accompanying noun. For now, I want to concentrate on two main points: the word *THAT* clearly has the same model of definiteness as *THE*, and it has an additional domain of meaning because its lexical entry includes the fact that the referent of *THAT* is (relatively) far from the speaker. Each utterance word has a speaker, and the lexical entry can model that fact; indeed, in the case of deictic expressions such as *THIS* and *THAT*, it is necessary for the lexical entry to show how the use of the word relates to the speech event. Crudely

speaking, the change in Old English *se* from demonstrative to determiner is the loss of the deictic element in its lexical entry. However, as we shall see, whatever semantics we adopt for definiteness, the story will be more complicated than just the loss of the deictic element.

Below, I discuss how the known facts can be modelled given the quantifier account of the and subsequently, I show how they could be modelled given the familiarity theory. However, before we move on to how the different theories model the change, we need to agree some ground rules: first, we need to establish what the most plausible syntactic account of demonstratives and definite articles will look like; and second we need to be clear about the main Old English facts, so that the models I propose can be evaluated in terms of their plausibility.

The syntax is straightforward, if a little non-standard. I am assuming that the determiners of English are all pronouns, and that they take common nouns as their complement. In this analysis, I am following Hudson (1990: 268-276). Note that the analysis is not the same as the Abney's version of NP structure, which makes the reverse categorial assumptions, that pronouns are determiners. The difference matters, because the pronouns-are-determiners analysis requires there to be a novel functional category, determiner, which has different distributional properties from the category noun, and which projects a novel functional node in a tree. In Hudson's WG story, the category "pronoun" is just a subtype of noun, so whether the noun selects the determiner or the other way around the NP is just headed by a noun.

We can see how the account works by comparing A(N), THE, THIS, THAT, and MY/MINE.

(26) a. *We saw a/the cat.*

b. **We saw a/the.*

(27) a. *We saw this/that/my cat.*

b. *We saw this/that/mine.*

The argument is simple: treat A(N), THE, THIS, THAT and MY/MINE as transitive pronouns. In (26a) and (27a) they occur with their complement; in (27b), THIS, THAT and MY/MINE occur without their complement. In (26b) we see that A(N) and THE are obligatorily transitive. As long as we treat MY and MINE as just conditioned variants of the same lexeme, where the different variant depends on whether the item has a complement or not, all of these items can be treated as members of the class of pronouns, which take complements. According to this story, “articles” and “demonstratives” are both subtypes of transitive pronouns; the difference between them is simply that an article must always have a complement and a demonstrative can have a complement or be elliptical or anaphoric.¹¹

In broad terms, the WG analysis of “determiners” as transitive pronouns makes the analysis of the emergence of the English definite article very simple: an earlier demonstrative loses its ability to occur without a complement (and to be anaphoric, or

interpreted elliptically) and has to occur at all times with its complement. At the same time, the ability of (singular count) nouns to occur without a determiner is also lost, giving rise to the mandatory appearance of determined NPs. The latter phenomenon is a reflex of the loss of case, and I shall not explore it here—if the job of a determiner is to fit up a noun to refer, then this is a job that nominal inflection also does. With the loss of case, comes the loss of being set up for reference.

Before I come to evaluating the different semantic arguments and how they accommodate the history of the English definite article, I briefly summarise the history of the English definite article. There are two main recent studies that I am aware of: Denison (2006) and Sommerer (2010). Denison's paper is a handbook article which addresses theoretical questions of how to identify word classes, and what their boundaries are; Sommerer's thesis is an investigation into the emergence of the English definite article from the Old English demonstrative *se*. Crucially, both of these studies assume that there is such a word class as "determiner", and so sets out to establish criteria for it as a word class, and to identify when this class might first be identified in the history of English.

Sommerer (2010) runs through a range of uses of the Old English demonstrative, as well as showing that in OE it was possible for bare nouns to occur in referring positions. The examples in (28) (Sommerer 2010: 14, example 1) show that Gothic, Old High German, Old English and Old Saxon could all have referring nouns which occurred without a demonstrative, or anything otherwise similar to an article.

- (30) *and þær ða burh getimbrede, _ þæs ilcan geares þa æt Bricge*
*and there the burgh built, and in the same year that at Bridgeworth*¹²

Example (31) is also anaphoric, but arguably to a clausal antecedent rather than a nominal one.

- (31) *ðaða he wæs gebroht to geleafan mid ðære grapunge,*
When he was brought to faith, with the touch,
þa wearð seo twynung þurh þæt us ætbroden
*then was the uncertainty by that from us taken*¹³

And the final example I shall give shows that the Old English demonstrative also functioned in relative clauses (of the correlative type, rather than clauses with the structure of modern relative clauses).

- (32) *Abel, Adames sunu, rihtwis and Gode andfenge,*
Abel, Adam's son, righteous and to God loyal,
þone ofsloh Cain his broðor
*whom/this-one slew Cain, his brother*¹⁴

Both Hawkins (2004) and Diessel (1999) observe that one possible grammaticalization path for relative markers is from demonstratives, so the final use is not surprising. This use of the OE demonstrative does not, however, last into Middle English. I take it that the loss of case reduces the functional utility of this construction as it can no longer serve to relativise low items on the Keenan-Comrie Accessibility Hierarchy.

The Old English facts are consistent with the account I have sketched where we treat both the Old English demonstrative and the modern English definite article as transitive pronouns, where the transitivity is optional in the case of the demonstrative. Which of the semantic approaches to definiteness might work best with this syntactic treatment, and what then would the diachronic story be?

First, whatever semantic story we might adopt, there will be no category emergence. Determiners will not, on my story, be a new category of English. What we see in the morphosyntax is a restriction of the definite article to a transitive structure—or in other terms the transitivity becomes obligatory. I will treat the narrowing of the structures the definite article can occur in—so that, unlike the demonstrative, it cannot occur in relative clauses—as a consequence of the more general loss of case.

I present the familiarity story first and then move to the set analysis. As demonstratives are deictic, if definite articles are familiarity expressions, then what must happen is that explicit reference to something in the context (whether the demonstrative has a complement or not) is replaced with reference to some entity which is understood by the speaker to be part of shared knowledge. So, if I say at home *the dog needs feeding*, my family will understand that Jumble, our pet dog, is the dog at

hand, not the mangy dog from the farm down the road which wanders around our neighbourhood.

How would this work? Deixis is subjective—it is anchored in the speaker and the speaker-centred situatedness of the utterance. In order to make such an analysis work, we would need to assume that familiarity was inter-subjective, because it is the expression of a meaning of which the hearer is an essential part. Why, if I did not expect there to be a dog familiar to my interlocutor, would I use the expression *the dog*? The expression creates a presupposition of familiarity for both speaker and hearer.

According to this story, there would be several ways of creating “definiteness”: an expression would be definite if it was deictic, because the referent could be contextually identified; it would be definite if it belonged in a shared knowledge of the world, because the speaker would be presupposing the hearer’s familiarity with the object; and so forth. Definiteness, then, would be a heterogeneous category, perhaps organised in a family resemblance structure.

On the other hand, this way of construing how definite expressions are definite requires two things: first, it requires an extension to Traugott’s (2010) notion of intersubjectivity. Traugott limits intersubjectivity to “addressee self-image”, but if definite expressions are intersubjective—as I have claimed a familiarity account must be—then the notion of intersubjectivity is extended to shared knowledge, or the presupposition of shared knowledge. Second, this account requires that there is a semantic change: as demonstratives become determiners, so they lose their deictic anchoring and acquire a new meaning: the presupposition of shared knowledge. In the

next two diagrams, I give representations of the demonstrative in the phrase *that man* and the definite article in the phrase *the man*, which show the relevant changes. These diagrams allow us to compare this model of the change with the alternative Russellian account of definite descriptions.

Figure 13 about here

In this diagram, and in the following one, I am treating both *that* and *the* as the head of the phrase. I am not complicating the diagrams by adding categorial information, but I am assuming that they are both pronouns. I am assuming that the only difference between THAT and THE, and by extension between THE and its demonstrative historical antecedent, SE, is in the semantics.

Figure 15 about here

As we can see from the diagrams, the difference between THAT and THE comes down to the loss of a particular meaning situated in the utterance event, and the acquisition of what Hudson calls the “directed identity” link (see the discussion of Figure 2 above). In both examples, the NP actually “refers” (in the intramental sense of reference) in that there is a referential concept that the words identify. According to this theory, definiteness is heterogeneous: there are several different semantics associated under

this single rubric. And according to this theory, the semantic change is as I described it above: from deictic to shared knowledge.

In the next diagram, I present a lexical entry for *se*, which adopts the semantics in Figures 13 and 14. The diachronic story, according to this theory, would be (i) that the semantics associated with the third sublexeme emerged, and then (ii) that the sublexemes split, to be associated with *THAT* and *THE* respectively.

Figure 15 about here

The lexeme *se* is where all of the inflectional information is collected. I have left this paradigmatic information out of the diagram in order to avoid unnecessary complexity.

The sublexeme *se*₁ is the use of *se* we find in example (29)—an anaphoric use of the pronoun. The sublexeme *se*₂ accounts for the demonstrative interpretation in example (30) and the final sublexeme, *se*₃ accounts for the definite, non-demonstrative, interpretation in that example. All of the representations are necessarily somewhat schematic and there is clearly more complexity: for example, I have shown that *se*₂ and *se*₃ have complements, but I have not shown that those complements are nouns, or that their referents are instances of their senses.

The diachronic story that Figure 15 shows is simple: the semantics of *se*₃ emerge, and then in time *se*₃ comes to be realised by *THE*; and *se*₁ and *se*₂ come to be realised by *THAT*₁ and *THAT*₂. *THE* and *THAT* come to be identified as separate lexemes, and the

inflectional information attached to the lexeme *SE* is lost. Apart from changing realisations, the main innovation is the semantic change from *SE2* to *SE3*.

On the set approach to the semantics, *that man* looks like Figure 16, and *the man* looks like Figure 17.

Figure 16 about here

The main difference between the analysis of *that man* and *the man* is simply the difference between a deictic expression and an expression that is not deictic.

Figure 17 about here

According to this theory, both *THE* and *THAT* express their definiteness in the same way: the difference between them is simply that *THAT* is deictic and *THE* is not. Now we are in a position to see what a lexical entry for *SE* would look like.

Figure 18 about here

As in the other analysis, *SE1* represents the sublexeme for the anaphoric pronoun in (29) and *SE2* the lexeme that must exist on the demonstrative interpretation of (28). *SE3* represents the definite article interpretation of the example in (28). As in the discussion of Figure 15, there is not much to the diachronic story: the semantics of *SE3* emerge, and

then SE3 comes to be realised by THE, and SE1 and SE2 come to be realised by that. The main difference is in the claim about the semantics of definiteness, and the semantic change. The diagram is as schematic as the previous one.

The diagram in Figure 18 make the claim that the only semantic change from transitive demonstrative to definite article was the loss of deictic meaning—nothing else happened. At the same time, the only change in form was the realisation of the definite article by THE, which is (as I noted above) a clitic—a change in form that has to be accommodated on either semantic theory. I show the subsequent stage in the relevant lexical entries for Modern English in Figure 19.

Figure 19 about here

In this diagram, I have shown that THE is a separate lexeme, and that SE1 and SE2 have come to be replaced by THAT1 and THAT2. I have also shown the morphosyntactic realizational facts, although I have not shown the two pronunciations of THE, which are contingent on whether the following word begins with a vowel or a consonant. I will not present a diagram for these words on the familiarity approach—it is this diagram, but with the semantics of Figure 15.

The diagrams show that either theory of definiteness can handle the grammaticalization of THE. I think, though, that we can do better than that, and consider how the theories might compare. Both stories—at least given the syntactic analysis I have adopted here—make the grammaticalization of THE simple. The common elements

are that there is a lexeme-split (SE to THAT and THE) and that THE evolves into a clitic. According to the account here, there is no categorial change. So the only difference between the two theories is that the familiarity story of definiteness requires us to assume that a new relation comes into play, which allows the speaker/hearer to identify a referential token in the (discourse) environment, whereas the set-theoretic account of definiteness merely involves the loss of the few network links that make the meaning of SE OR THAT deictic.

The familiarity approach requires an extension of the notion of intersubjectivity, whereas the set-based approach requires no such extension of the analytical tools. And finally, the familiarity approach requires us to treat definiteness as a heterogeneous category, whereas the set-based approach allows us to treat it as a unitary phenomenon, with the differences between the different meanings being due to such additional network links as are necessary for an account of the different meanings of the various definite expressions listed in §2 (at least with the exception of proper names). Given the advantages for linguistic analysis of treating definite expressions as proportional quantifiers that I pointed out in §7, and given the far simpler story of grammaticalization that it gives us, I conclude that functional linguistics working on grammaticalization, and cognitive linguists working on the phenomena discuss in §7, will want to work with a treatment of definite expressions as proportional quantifiers.

9. Conclusion

I have shown how it is possible to model the treatment of definite articles as proportional quantifiers in at least one cognitive theory—a theory which, by taking language to be a classified symbolic network also makes it fairly straightforward to treat variation and change by modelling how networks vary and change. Word Grammar is fully compatible with cognitive and functionalist assumptions about language—those are my own assumptions, after all—and it also allows us a full and precise modelling of the necessary linguistic structures and of how language changes. There is, in addition, one further point to make: cognitive and functional theories are interested most of all in language **use**. I hope that in this paper I have shown how formal findings can be integrated into a cognitive theory, and how in modelling the diachrony that is widely reported to do with the emergence of definite articles, we have also seen something novel about language use and how it plays out in language change. I have also shown how research findings from the formal literature can be implemented in a cognitive model, and even evaluated in a cognitive model: my treatment of the grammaticalization of **THE** is not just an analysis of the history, but is also an evaluation of the claim that **THE** is best treated as a quantifier.

References

- Abbott, Barbara. 2004. Definiteness and indefiniteness. In Laurence R. Horn & Gregory Ward., eds., *Handbook of Pragmatics*, 122-149. Oxford: Blackwell.

- Abney, Steven. 1987. The English noun phrase in its sentential aspect. PhD dissertation, MIT, Cambridge, MA.
- Barwise, Jon and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4 /2. 159--219.
- Bolinger, Dwight. 1977. There. In *Meaning and form*, 90-123. London: Longman.
- Burton-Roberts, Noel. 1989. *The limits to debate: a revised theory of semantic presupposition*. Cambridge: CUP.
- Cann, Ronnie, Ruth Kempson & Lutz Marten. 2005. *The dynamics of language*. Oxford, Elsevier.
- Chesterman, Andrew. 1991. *On definiteness: a study with special reference to English and Finnish*. Cambridge: CUP.
- Christopherson, Paul. 1939. *The articles: a study of their theory and use in English*. Copenhagen: Munksgaard.
- Clark, Herbert H. & Catherine R. Marshall. 1981. Definite reference and mutual knowledge. In Avarind K. Joshi, Bonnie L. Webber & Ivan A. Sag, eds., *Elements of discourse understanding*, 10-63. Cambridge: Cambridge University Press.
- Davidse, Kristin. 2004. The interaction of quantification and identification in English determiners. In Michel Achard and Suzanne Kemmer (eds.) *Language, culture and mind*, 507-533. Stanford: CSLI Publications..
- Declerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudo-clefts*. Leuven: Leuven University Press.

- Denison, David. 2006. Category change and gradience in the determiner system. *The handbook of the history of English*, 289-304. Ans van Kemenade & Bettelou Los (eds.). Oxford: Basil Blackwell.
- Diessel, Holger. 1999. *Demonstratives: form, function and grammaticalization*. Amsterdam: Benjamins.
- Donnellan, Keith. 1966. Reference and definite descriptions. *The Philosophical Review* 75/3: 281-304.
- Elbourne, Paul. 2010. The existence entailments of definite descriptions. *Linguistics and Philosophy* 33(1): 1–10.
- Fauconnier, Gilles. 1994. *Mental spaces*. Cambridge: CUP.
- Gazdar, Gerald. 1979. *Pragmatics: implicature, presupposition and logical form*, New York: Academic Press.
- Gisborne, Nikolas. 2010. *The event structure of perception verbs*. Oxford: OUP.
- Gisborne, Nikolas and Amanda Patten. 2011. Constructions and grammaticalization. In: *The Oxford handbook of grammaticalization* 92-104. Berndt Heine and Heiko Narrog, (eds.), Oxford: OUP
- Hawkins, John A. 1978. *Definiteness and indefiniteness*. London: Croom Helm.
- Hawkins, John A. 2004 *Efficiency and complexity in grammar*. Oxford: OUP
- Heim, Irene. 1983. File change semantics and the familiarity theory of definiteness. In: Rainer Bäuerle, Christoph Schwarze and Arnim von Stechow (eds), *Meaning, Use and Interpretation* 164-90. Berlin: Walter de Gruyter.
- Heycock, Caroline and Anthony Kroch. 1999. Pseudocleft connectivity: implications for the LF Interface. *Linguistic Inquiry* 30.3: 365-397

- Hopper, Paul & Elizabeth Traugott. 2003. *Grammaticalization*. 2nd edn. Cambridge: CUP.
- Hudson, Richard. 1984. *Word Grammar*. Oxford: Blackwell.
- Hudson, Richard. 1990. *English word grammar*. Oxford: Blackwell.
- Hudson, Richard. 2007. *Language Networks: the new Word Grammar*. Oxford: OUP.
- Hudson, Richard. 2010. *An Introduction to Word Grammar*. Cambridge: CUP.
- Jackendoff, Ray. 2002. *Foundations of language*. Oxford: OUP
- Kamp, Hans. 1981. A theory of truth and semantic representation. In Jeroen Groenendijk, *et al* (eds) *Formal methods in the study of language* (Amsterdam: Mathematical Centre) 277-322. [Reprinted in J. Groenendijk, T. Janssen and M. Stockhof (eds) (1984) *Truth, interpretation and information*. Dordrecht: Floris. 1-41.]
- Karttunen, Lauri. 1976. Discourse referents. In J. McCawley (ed.) *Syntax and Semantics vol. 7: notes from the linguistics underground*. New York: Academic Press, 363-385.
- Kearns, Kate. 2000. *Semantics*. Basingstoke: Palgrave.
- Kempson, Ruth. 1975. *Presupposition* and the delimitation of semantics. Cambridge: CUP.
- Kuno, Susumu. 1972. Functional sentence perspective: a case study from Japanese and English. *Linguistic Inquiry* 3, 269-320.
- Lakoff, George. 1987. *Women, fire and dangerous things*. Chicago: Chicago University Press.
- Langacker, Ronald. 1991. *Foundations of cognitive grammar*, Volume II, Descriptive Application. Stanford, CA: Stanford University Press.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: CUP.

- Mikkelsen, Line. 2005. *Copular clauses: specification, predication and equation*.
Amsterdam: John Benjamins.
- Milsark, Gary. 1977. Toward an explanation of certain peculiarities of the existential construction in English. *Linguistic Analysis* 3:1-29.
- Mitchell, Bruce. 1985. *Old English syntax* (2 volumes) Oxford: Clarendon Press.
- Moro, Andrea. 1997. *The raising of predicates. Predicative noun phrases and the theory of clause structure*. Cambridge: CUP.
- Neale, Stephen. 1990. *Descriptions*. Cambridge, MA: MIT Press.
- Patten, Amanda L. 2010. Clefts, construction grammar, and grammaticalization.
University of Edinburgh PhD dissertation.
- Prince, Ellen F. 1992. The ZPG letter: Subjects, definiteness, and information status. In W. C. Mann & S. A. Thompson, eds., *Discourse description: Diverse linguistic analyses of a fund-raising text* 295-326. Philadelphia: John Benjamins.
- Russell, Bertrand. 1905. On denoting. *Mind* 14, 479-493
- Sommerer, Lotte. 2010. Old English *se*: from demonstrative to article. A usage-based study of nominal determination and category emergence. PhD dissertation,
University of Vienna.
- Strawson, Peter F. 1950. On referring. *Mind* 59, 320-344
- Traugott, Elizabeth. 1982. From propositional to textual and expressive meanings: some semantic-pragmatic aspects of grammaticalization. In Lehmann, Wilfried. P. & Yakov Malkiel (eds.), *Perspectives on Historical Linguistics* 245-271. Amsterdam: Benjamins,.

- Traugott, Elizabeth. 1992. Syntax. In Hogg, R. M. (ed.). *The Cambridge history of the English language. Volume 1. Old English. From the beginnings to 1066*. Cambridge: Cambridge University Press, 168-289.
- Traugott, Elizabeth. 2010. Intersubjectivity, and (inter)subjectification: a reassessment. In Kristin Davidse, Lieven Vandelanotte, & Hubert Cuyckens (eds.) *Subjectification, intersubjectification and grammaticalization*. Berlin: Mouton de Gruyter, pp. 29-71.
- Van Langendonck, Willy. 1979. Definiteness as an unmarked category. *Linguistische Berichte* 63, 33-55.
- Van Langendonck, W. 1994. Determiners as heads? *Cognitive Linguistics* 5, 243-260.
- Wiggins, David. 1980. "Most" and "All": some comments on a familiar programme, and on the logical Form of quantified sentences. In M. Platts (ed.) *Reference, truth and reality*, 318-346. London: Routledge and Keegan Paul.
- Wilson, Deirdre. 1975. *Presupposition and non-truth-conditional semantics*. New York: Academic Press.

Diagrams

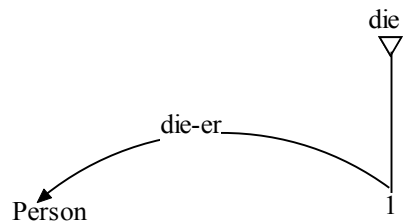


Figure 1. 'Everybody dies' from Hudson (2007: 31)

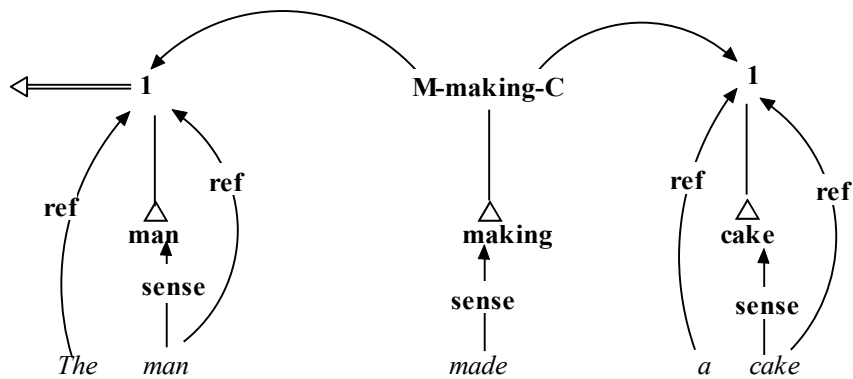
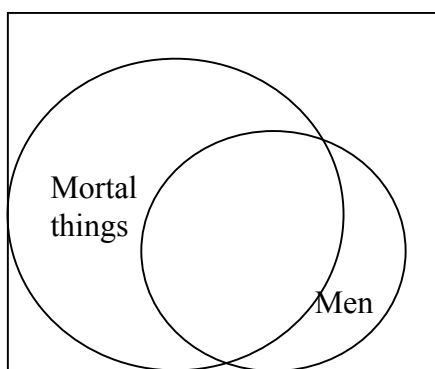


Figure 2. Hudson's treatment of definiteness



Most men are mortal is true iff the greater part of the set of men intersects with the set of mortal things.

Figure 3. A set interpretation of *most men are mortal*

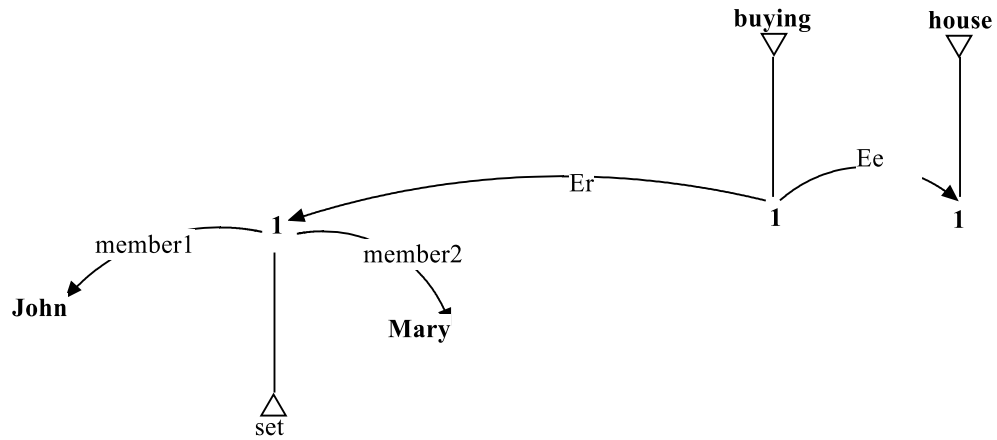


Figure 4. Collective interpretation of *John and Mary bought a house*

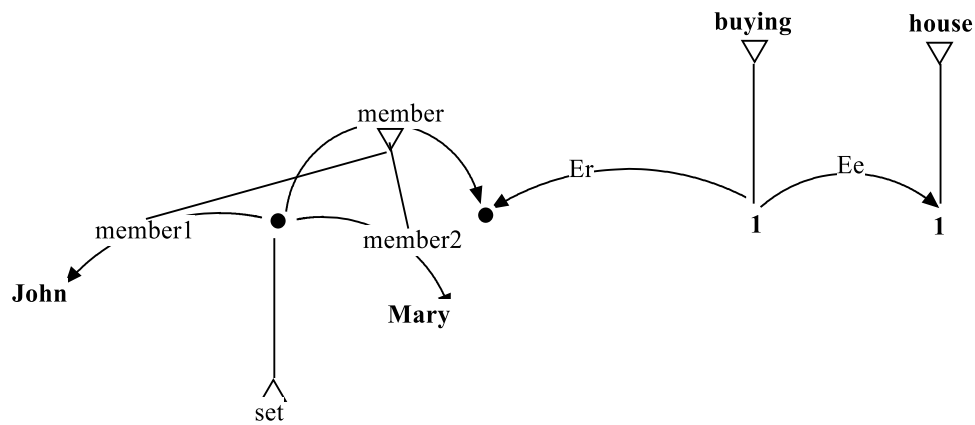


Figure 5. Distributive interpretation of *John and Mary bought a house*

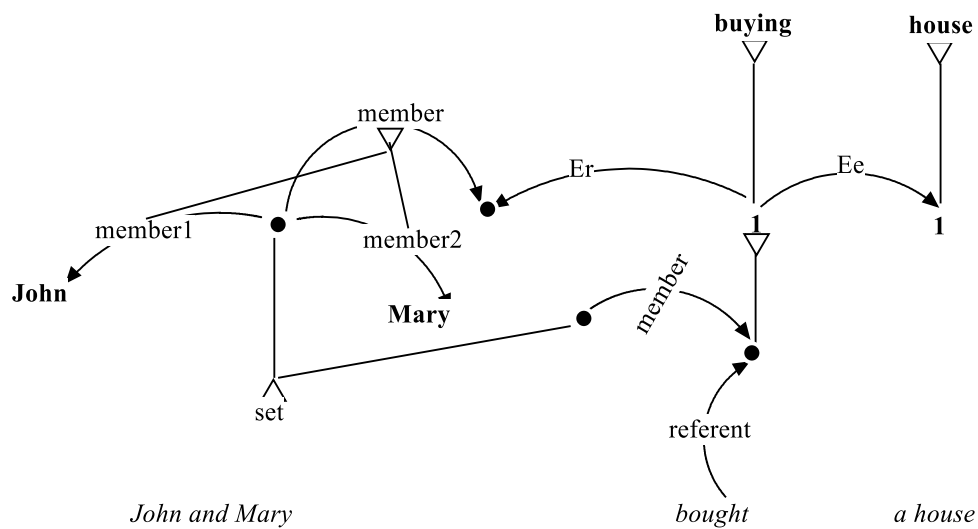


Figure 6. Distributed interpretation of *John and Mary bought a house* with plural events.

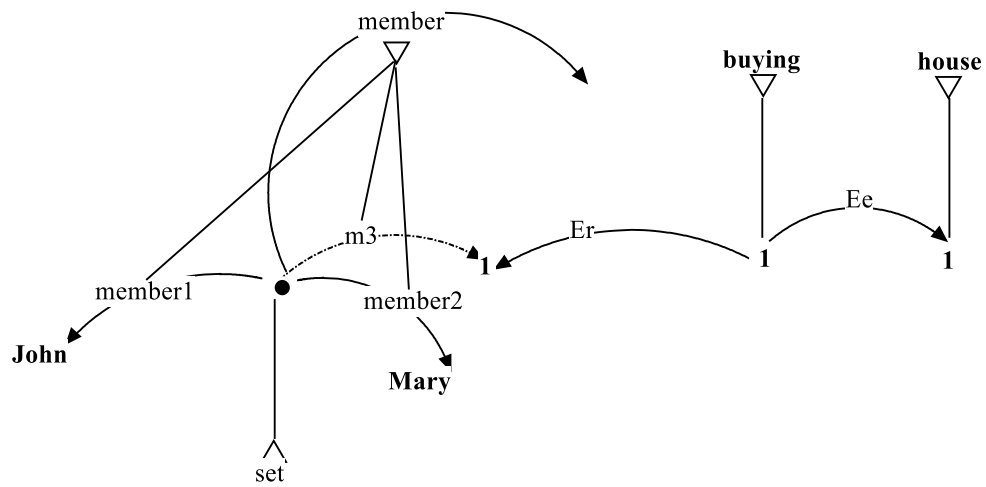


Figure 7. John or Mary bought a house

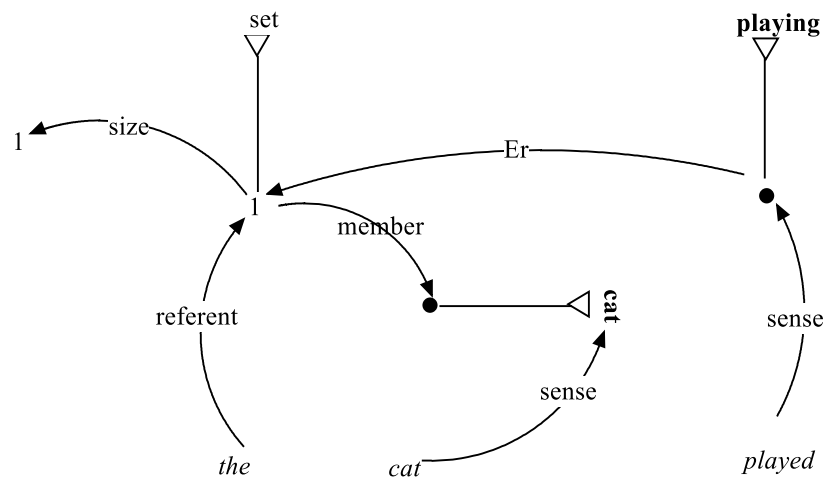


Figure 8. THE is a quantifier, singular NP

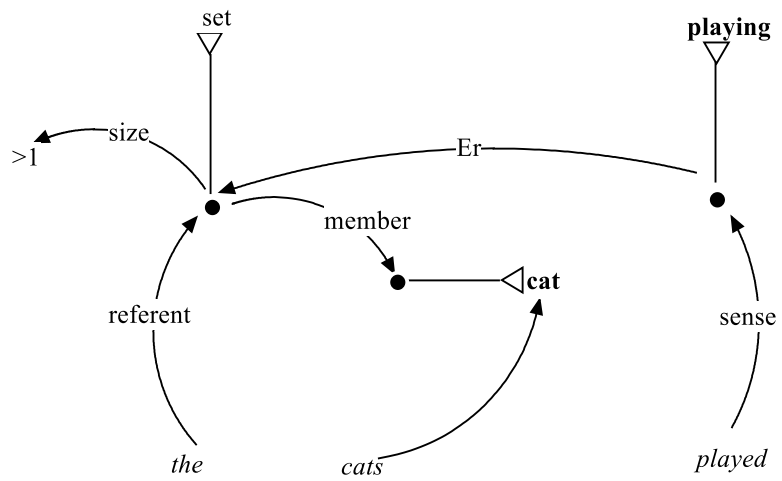


Figure 9 THE is a quantifier, plural NP

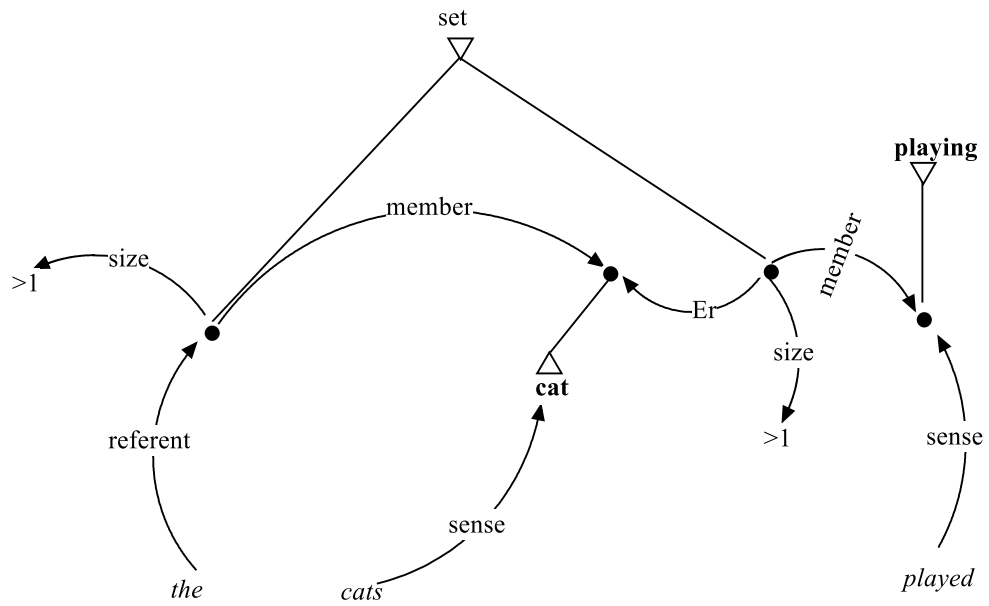


Figure 10. Distributive interpretation of a definite NP

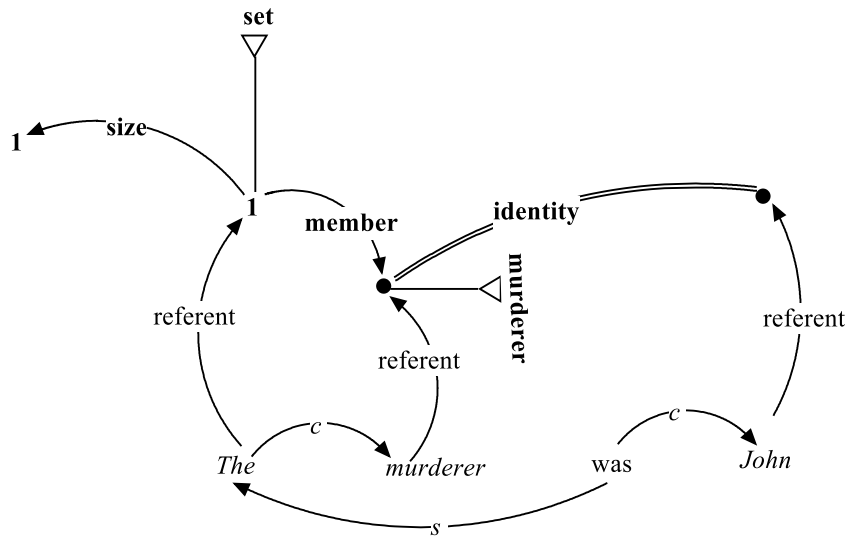


Figure 11. *The murderer was John*

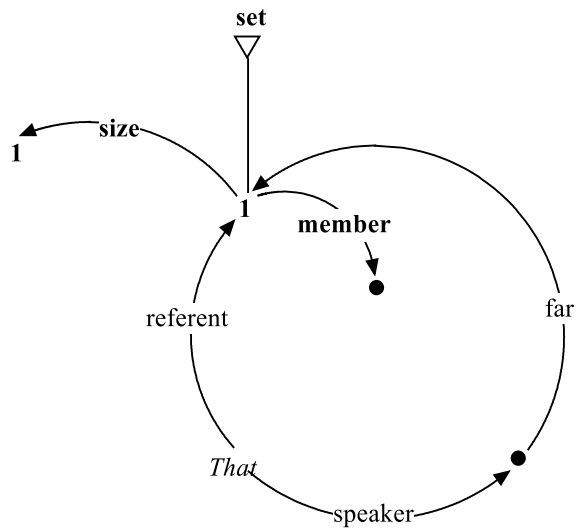


Figure 12. A partial WG lexical entry for THAT.

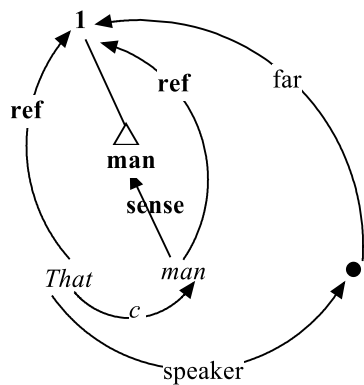


Figure 13. *That man* in WG, without sets

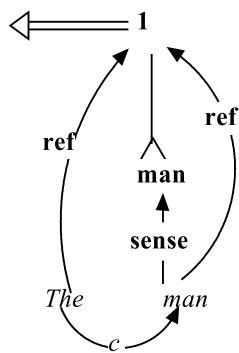


Figure 14. *The man* in WG, without sets

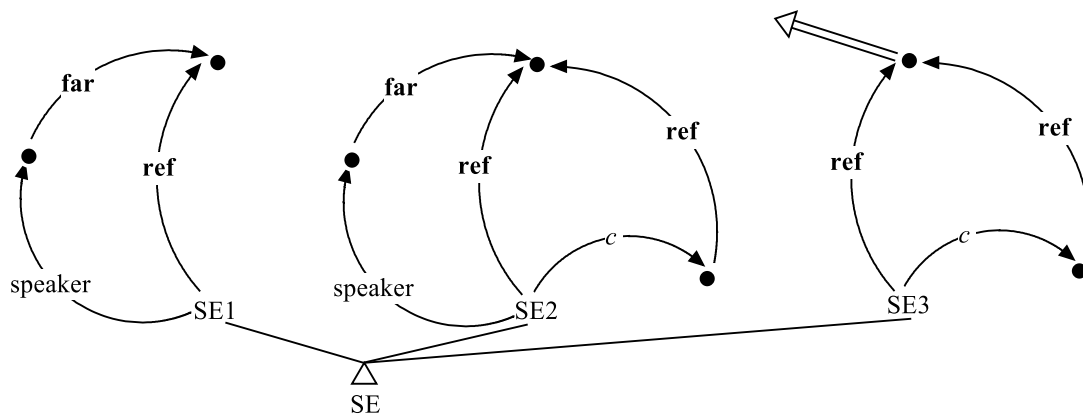


Figure 15. A lexical entry for *se*, on the familiarity story.

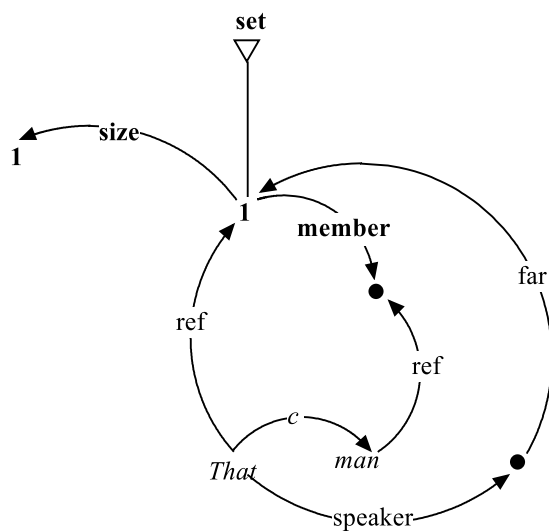


Figure 16. *That man* in WG, with set-theoretic definiteness

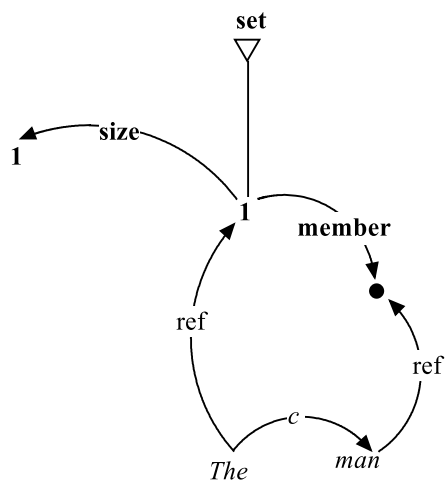


Figure 17. *The man* in WG, with set-theoretic definiteness

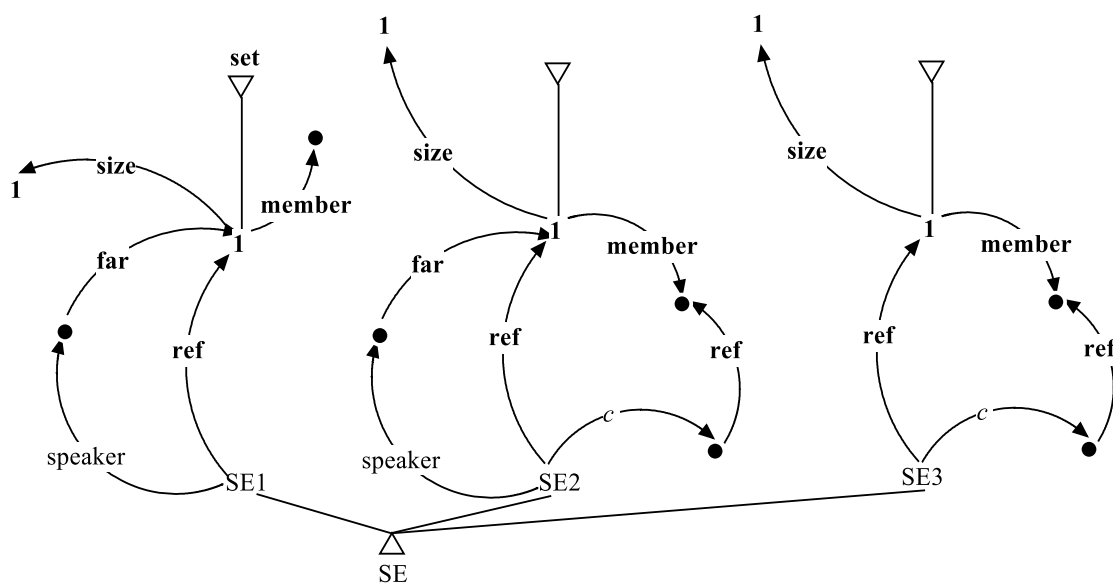


Figure 18. The lexeme *SE*, with a set-based semantics.

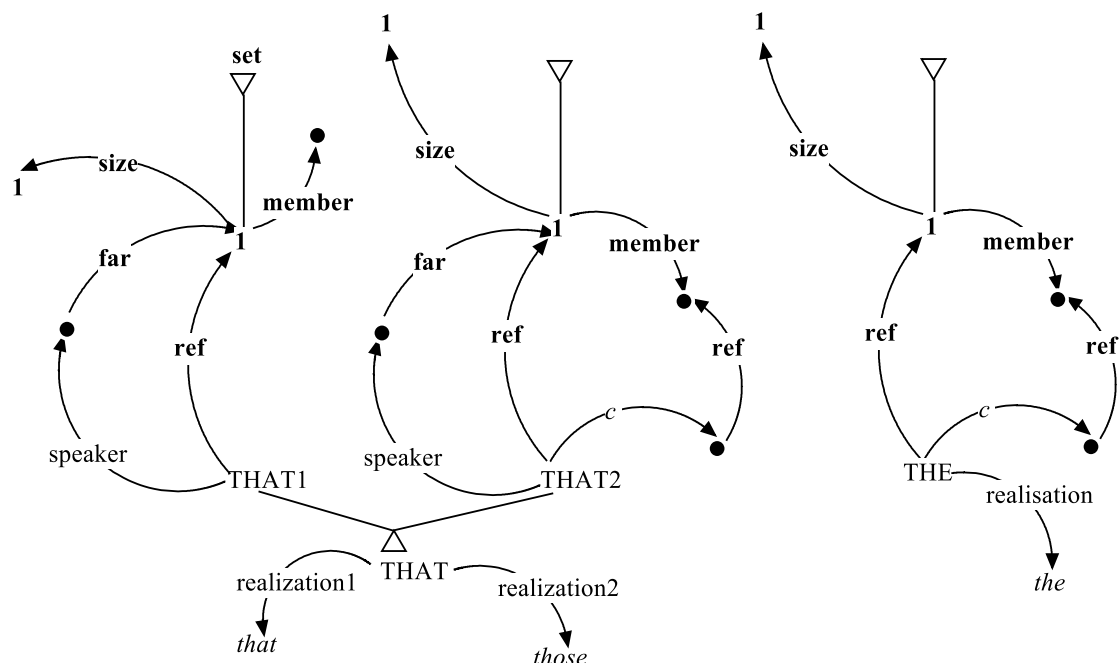


Figure 19. Lexical entries for Modern English THAT and THE.

* I would like to thank Willem Hollmann, Ekkehard Koenig and an anonymous referee for a number of suggestions improving the paper. I should also like to thank the audience at the Societas Linguistica Europaea Conference in Vilnius, Ronnie Cann and Geoff Pullum for useful comments and advice, and Dick Hudson for reading the paper through and making extensive comments.

¹ Langacker (1991: 92, 98) does not actually say explicitly that he views THE as a proportional quantifier, although his semantic glosses are compatible with such a view. Langacker (1991: 98) says, “the designated instance t_i of T is unique and maximal in relation to the current discourse space.”

² Although, of course, THERE can be stressed in the first example, and so receive a locative interpretation as well.

³ Although note that MANY at least can also have a cardinal interpretation (Milsark 1977).

⁴ Neale (1990: 41-2) argues whether the appropriate treatment is restricted quantification or binary quantification. He observes that the two systems have the same expressive power, and asserts that as restricted quantification most closely matches natural language syntax, it is the system he will prefer.

⁵ Kearns’ term “strong quantifier” simply means “proportional quantifier”.

⁶ The discussions of (11) and (12) are foregrounded to some extent by Hawkins (1978: 157-167), where the notion of uniqueness is replaced by one of inclusivity in order to capture the semantics of definite plurals. Van Langendonck (1979) challenges the notion of inclusivity on the grounds that plurals can be used without all members of the set being invoked, as in *the clouds were covering the moon* (see also Chesterman 1991: 22-24 for discussion). I think that Van Langendonck’s problem can be resolved by thinking about definiteness as existential and universal quantification over a restricted set with reference to a background set.

⁷ The semantic relations Er and Ee are glosses over the many semantic relations that verb meanings might have. The Er is the relation that relates to the referent of the Subject in the active voice; the Ee links to the referent of the Object, when the verb is transitive.

⁸ The picture is a little more complex in that when SOME has a schwa rather than a full vowel, it does not imply a superset, and it can also go together with mass nouns, giving rise to a part-of relation. Worse, it

can also be used non-proportionally when set off against another cardinal value: *there were some – not many – misprints*. I am grateful to an anonymous reviewer for pointing this out.

⁹ It is necessary to make sure that we agree on which subtype of the existential THERE construction is under discussion. There are apparent counter-examples to the generalization: *there was the most beautiful sunset I've ever seen* and *there was the odd drop of rain*. As Kearns (2000: 81-85) points out, there are THERE+BE constructions which permit definites: the main diagnostic sentences are the basic existential THERE+BE sentences.

¹⁰ Declerck (1988: 47ff) offers a different, but semantic, analysis of specificational sentences: he argues that they express a variable-value relationship which exploits Donnellan's (1966) treatment of definite expressions, with there being two uses: an attributive use and a referential one. Declerck exploits this distinction to capture the asymmetry between the two NPs; his analysis is a variant of the concealed question approach, and essentially takes the semantics of specificational questions to be the answer to an unspoken question. However, as Patten (2010: 65-66) points out, the subjects of specificational sentence cannot be treated as either strongly or weakly referring and so cannot be attributive. Patten's evidence is that weakly referring NPs (as in Donnellan's 1966 example *Smith's murderer is insane*) pronominalize with HE/SHE, whereas specificational subjects pronominalize with IT.

¹¹ Hudson (2004), following a paper by Van Langendonck (1994) revisits his earlier analysis, and argues that determiners (or transitive pronouns) are categorically as he described them in (1990), but that there is also a mutual dependency relation between them and their noun complements. The argument does not affect the analysis of the category of demonstratives and articles, so there is no need to visit all of the claims in the paper. Van Langendonck's arguments are mainly that the distribution of NPs follows from the noun and not the determiner. This, then, makes the noun the head. Hudson (2004) effectively splits headedness into two kinds: distributional headedness, and structural headedness (see also Rosta 2005). He argues that the determiner is the structural head—i.e. the head that determines the internal structure of the NP—and the noun is (or can be) the distributional head. For my purposes here, this two-headed analysis works. Determiners are still classified as (transitive) pronouns, and the distributional facts that Van Langendonck identifies can still be accommodated.

¹² (ChronC 96.31 (912))

¹³ Sommerer (2010: 21) example (6); Ælfric's Homilies i. 234.23.

¹⁴ Sommerer (2010: 22) example (11); Ælfric's Homilies ii.58.25.